How AMC (Arthrogryposis) and CP (Cerebral Palsy) have to adjust to the “new genetics” and Epigenetics

Judith G. Hall, OC, MD
Departments of Pediatrics and Medical Genetics
University of British Columbia
Vancouver, BC, Canada
- No known conflicts of interest -

PLAN OF TALK
• Thank you for inviting me
• Clinical Geneticist
• Long standing interest in arthrogryposis (AMC)
• Common challenges
• Non-traditional inheritance
• Epigenetics

ARTHROGRYPOSIS
(MULTIPLE CONGENITAL CONTRACTURES)
Congenital non-progressive limitation of movement of two or more joints in different body areas (lots and lots of things get included)

WHY SO INTERESTING?
• All about embryo/fetus developing movement
• All causes share decreased in utero movement
• Why develop contractures?

CONGENITAL CONTRACTURES IN THE NEWBORN
• Clubfoot...........................................1/500 – 1/1000
• Congenital dislocated hips.............1/200 – 1/500
• Multiple congenital contractures.....1/3000 – 1/6000
• All congenital contractures..........1/100 – 1/250
Arthrogryposis is not a diagnosis — it is a sign

LIMITATION OF FETAL JOINT MOBILITY
MULTIPLE CONGENITAL CONTRACTURES (ARTHROGRYPOSIS)

FREQUENCY OF ARTHROGRYPOSIS
- Many lethal types, miscarriages, stillborns
  - VERY heterogeneous
  - No proper ICD code(s)
  - Need population base information
  - Australia epidemic
  - North America
    - Washington State
    - British Columbia registry
  - Others
    - Finland
    - Sweden

OCCURRENCE OF ARTHROGRYPOSIS
- ~ 1/3000 – 5000...live births
- 1/3..................Amyoplasia
- 1/3..................CNS – newborn lethal
- 1-3..................Heterogeneous group of disorders

WAYS TO APPROACH A DIAGNOSIS –
What are useful clinical discriminators?

AREAS OF INVOLVEMENT TOTAL STUDY GROUP

<table>
<thead>
<tr>
<th>Primary limbs</th>
<th>Limbs plus other body areas</th>
<th>Limbs plus CNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>166</td>
<td>129</td>
<td>33</td>
</tr>
<tr>
<td>53%</td>
<td>37%</td>
<td>10%</td>
</tr>
</tbody>
</table>

AMYOPLASIA “CLASSICAL ARTHROGRYPOSIS”
- Typical symmetric positions of limbs
- Usually “teratogenic” clubfoot
- Absent muscles with fibrotic replacement
- Mid facial hemangioma
- 10% abdominal structural anomaly (vascular accident) and other vascular compromise (lost fingers or toes)
- Apparent increase in one of monozygotic twins
- Surprisingly good response to early physical therapy
- No apparent recurrence risk or risk for other congenital anomalies
AMYOPLASIA

“CLASSICAL ARTHROGYROSIS”

• Typical symmetric positions of limbs
• Usually “teratogenic” clubfoot
• Absent muscles with fibrotic replacement
• Mid facial hemangioma
• 10% abdominal structural anomaly (vascular accident) and other vascular compromise (lost fingers or toes)
• Apparent increase in one of monozygotic twins
• Surprisingly good response to early physical therapy
• No apparent recurrence risk or risk for other congenital anomalies

AMYOPLASIA

— UPPER ONLY

AMYOPLASIA

— LOWER ONLY

AMYOPLASIA

and NORMAL TWIN
AMYOPLASIA
“CLASSICAL ARTHROGRYPOSIS”

- Typical symmetric positions of limbs
- Usually “teratogenic” clubfoot
- Absent muscles with fibrotic replacement
- Mid facial hemangioma
- 10% abdominal structural anomaly (vascular accident) and other vascular compromise (lost fingers or toes)
- Apparent increase in one of monozygotic twins
- Surprisingly good response to early physical therapy
- No apparent recurrence risk or risk for other congenital anomalies

PROPOSED PATHOGENESIS

CLASSIFICATION OF DISTAL AMCs

CLASSIFICATION OF DISTAL AMCs

Approach to multiple congenital contractures - clinical

- Mainly limbs
- Limbs and other body areas
- Limbs and CNS/lethal
<table>
<thead>
<tr>
<th>Hall</th>
<th>Distal</th>
<th>Handled</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Distal</td>
<td>I</td>
<td>TPM2</td>
</tr>
<tr>
<td>II</td>
<td>Gordon [cleft palate, 10]</td>
<td>I</td>
<td>MYH3</td>
</tr>
<tr>
<td>III</td>
<td>Ophthalmoplegia [fine muscle]</td>
<td>I</td>
<td>TNNT3, TNN12</td>
</tr>
<tr>
<td>IV</td>
<td>Syndactyly 1</td>
<td>I</td>
<td>MYH3</td>
</tr>
<tr>
<td>V</td>
<td>Trisomy 18, absent Hands + KB</td>
<td>I</td>
<td>MYH3</td>
</tr>
<tr>
<td>Freeman-Sheldon syndrome</td>
<td></td>
<td>V</td>
<td>MYH3</td>
</tr>
<tr>
<td>Shokouh et al.</td>
<td></td>
<td>II</td>
<td>TPM2</td>
</tr>
<tr>
<td>Shokouh et al.</td>
<td></td>
<td>II</td>
<td>TPM2</td>
</tr>
<tr>
<td>Shokouh et al.</td>
<td></td>
<td>II</td>
<td>TPM2</td>
</tr>
<tr>
<td>Chitayat, AR, DD</td>
<td></td>
<td>X</td>
<td>11q25</td>
</tr>
<tr>
<td>Chitayat, AR, DD</td>
<td></td>
<td>X</td>
<td>MYH8</td>
</tr>
<tr>
<td>Chitayat, AR, DD</td>
<td></td>
<td>X</td>
<td>FBN2</td>
</tr>
</tbody>
</table>
FETAL AKINESIA DEFORMATION SEQUENCE
PENA SHOKIER PHENOTYPE

- Intrauterine growth restriction
- Congenital contractures of the limbs
- Hypoplastic lungs
- Short umbilical cord

- Polyhydramnios – short gut
- Craniofacial anomalies – Micrognathia +/- small mouth
- +/- cleft palate
- High bridge of nose
- Depressed tip of nose

Lack of normal mechanical forces may lead to secondary deformations

“Use” is essential for normal development

SECONDARY EFFECTS FROM LACK OF MOVEMENT IN UTERO

- IUGR – limbs are short
- Contractures with “collagenosis”, extra connective tissue, thick capsule
- Abnormal relationship of limb to weight bearing – joints at odd angles
- Muscle – disuse atrophy, decreased mass
- Dimples – attached to overlying skin
- Other changes of FADS – lungs, gut, craniofacial, etc.

GROWTH

- AMC affected limbs - short and small
- Final height ~ 5th centile for family
- Less muscle and less calcification of bone means less weight
- Avoid obesity – makes for more work
- Some limbs grow even less normally (like post-polio)

As organs begin to function muscles begin to contract stretching developing tissues from inside and outside
EMBRYONIC/FETAL MOVEMENT

<table>
<thead>
<tr>
<th>Week</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Heart beating begins</td>
</tr>
<tr>
<td>5-6</td>
<td>Head and trunk “stirs”</td>
</tr>
<tr>
<td>7</td>
<td>Shoulders “shrug”</td>
</tr>
<tr>
<td>8</td>
<td>Rhythmic “breathing” begins even though larynx not open</td>
</tr>
<tr>
<td>9</td>
<td>Jaw starts to move</td>
</tr>
<tr>
<td>10</td>
<td>Upper arms moving</td>
</tr>
<tr>
<td>11</td>
<td>Hips, lower arms moving</td>
</tr>
<tr>
<td>12</td>
<td>Lower limbs kicking</td>
</tr>
<tr>
<td></td>
<td>Hands open and ankles moving into correct position</td>
</tr>
</tbody>
</table>

CLUBFOOT 16 WEEKS

Ultrasound 16 weeks
Clubfoot diagnosis
Amyoplasia
RS 517518 CH&MC
LTS3-85

PRENATAL DIAGNOSIS BY ULTRASOUND
(WHAT ARE THE CLUES, WHAT TO LOOK FOR)

- Usually not picked up without long careful real-time US study – 45 min – 1 hr
- Nuchal edema
- Thin undercalcified bones
- Movement may start any time from 11 weeks to 34 weeks
- Small lungs
- Diaphragm defect or decreased movements
- Other structure or space constraints (amniotic bands, uterine fibroid, amount of amniotic fluid)

2 RETROSPECTIVE STUDIES OF AMYOPLASIA

2013 and 2015

Only 25% of 4 limb Amyoplasia diagnosed prenatally!

RECURRENT RISK

<table>
<thead>
<tr>
<th></th>
<th>Primary limbs</th>
<th>Limbs plus other areas</th>
<th>Limbs plus CNS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated RR for overall group</td>
<td>Parents: 6.0%</td>
<td>6.5%</td>
<td>10.8%</td>
<td>10.4%</td>
</tr>
<tr>
<td>Estimated RR when knowns excluded</td>
<td>Parents: 4.7%</td>
<td>1.4%</td>
<td>7%</td>
<td>5%</td>
</tr>
</tbody>
</table>

If, and only if, a specific diagnosis cannot be made should a 5% recurrence risk estimation be given

Prognosis depends on the specific diagnosis and the natural history of that disorder
WHAT DOES THIS HAVE TO DO WITH CEREBRAL PALSY?

<table>
<thead>
<tr>
<th>AMC</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence</td>
<td>1/3000</td>
</tr>
<tr>
<td>Variability</td>
<td>&gt; 400 types</td>
</tr>
<tr>
<td>Seizures/PX</td>
<td>1/3</td>
</tr>
<tr>
<td>Prenatal Dx</td>
<td>Check movement (only at 2/3)</td>
</tr>
<tr>
<td>NRA</td>
<td>MRI, IUGR, Infection of decreased muscle</td>
</tr>
<tr>
<td>PT/OT</td>
<td>Essential mobilization</td>
</tr>
<tr>
<td>Drugs/Help</td>
<td>No</td>
</tr>
</tbody>
</table>

Both AMC (CNS group = 1/3) and CP (2/3) with DD and/or seizures...

- Look for chromosomal deletion
- Specific gene(s) mutations
GENE ONTOLOGY

- AMC – 330 genes placed in 27 pathways
- Find pathway of genes in development
- Possible therapeutic approach


ESCOBAR SYNDROME
(ONE OF MANY MULTIPLE PTERIGIUM SYNDROMES)

- Deficiency of one of the three embryonic components of the neuroreceptor
- 3 adult neuroreceptor components genes are normal
- Natural history – recurrent respiratory infections die in 20s

ESCOBAR SYNDROME
(ONE OF MANY MULTIPLE PTERIGIUM SYNDROMES)

- It is possible to “wake up” the adult receptor and “cure” the syndrome
- Therapy already available
- Tensilon-like drugs used in Myasthenia Gravis

So if you know the mechanism, you may be able to treat!

DEVELOPMENTAL BIOLOGY

- Time specific, tissue specific, gender specific gene expression
- Cascades, pathways, and networks – FGF, SHH, WNT, TGFβ, RTK-RAS
- Embryo, early fetus, viable fetus

STAGES OF AN INDIVIDUAL

1. Preimplantation
2. Embryo
3. Early fetus
4. Later fetus (adolesce)
5. Neonate
6. Infant
7. Young child
8. Older child
9. Adolescence (gender differences)
10. Adult (premenopause/postmenopause)
   A. Young
   B. Middle
   C. Mid-90s
11. Old
12. Old Old

RECENT WORK

- 22,000 genes – for 200,000 proteins
- Less than 2% of the genome contains protein coding genes
- Non-coding DNA highly conserved
- Commonality of animal species (90% mice, 98% primates, 99.9% humans)
- At least 20% of the human genome has copy number variation (CNV) between individuals (which reflect how individual we are)
- At least 100 “mutations” not in parents – 5 of which are “of unknown significance”
- There is no such thing as a “normal” human being
RECENT GENETIC WORK

• Genes are part of pathways/networks/systems (i.e., housekeeping, skin, limb, etc.)
• Maybe 1000 pathways – used over & over with alternative splicing/slightly different proteins
• Crosslink, buffering, redundancy of pathways
• Interactions with environment sculpt the pathways

NON-TRADITIONAL MECHANISMS (revealed by studying congenital anomalies)

• Mosaicism
• Microchimerism
• Imprinting
• Epigenetics
• Microbiome
• Fetal programming

MOSAICISM
• An individual comprised of two or more genetically distinct populations of cells derived from a single zygote (conception)

CHIMERA
• An individual who is a mixture of cells from two or more genetically distinct populations of cells which have been derived from more than one zygote

MICROCHIMERISM
• The presence of a small number of cells derived from another individual (transfusion, transplantation, pregnancy, twin, etc.)

McCUNE ALBRIGHT SYNDROME

• Mosaic for mutation causing increased activity of Gs protein
• Many different mutations, but consistent within an individual
• Would be lethal if present in all cells
• The affected individual is rescued by “normal” cells

MOSAICISM OCCURS IN ALL HUMAN BEINGS

• Large multicellular organism
• $10^{15} - 10^{14}$ cells in our bodies lost (plus programmed cell death)
• Mutation rate is 1/6,000 – 1/60,000 for known single gene human disorders
• We are all a walking “bag” of mutations
WHEN AND WHERE DOES/DID THE MOSAICISM ARISE?
A. Early – all tissue
B. Placenta – only (trisomy rescue)
C. Somatic – frequent
D. Germline (anticipation, multiple affected offspring) in normal appearing parent

NON-TRADITIONAL MECHANISMS (revealed by studying congenital anomalies)
- Mosaicism
- Microchimerism
- Imprinting
- Epigenetics
- Microbiome
- Fetal programming

EFFORTS AT NON-INVASIVE PRENATAL DIAGNOSIS
- Can find fetal cells in mother’s blood
- However, they are from all her pregnancies
- Somehow these cells are tolerated long term

FETAL MATERNAL MICROCHIMERISM
CHIMERA
- An individual who is a mixture of cells from two or more genetically distinct populations of cells which have been derived from more than one zygote
MICROCHIMERISM
- The presence of a small number of cells derived from another individual (transfusion, transplantation, twin, etc.)
MOASIC
- An individual comprised of two or more genetically distinct populations of cells derived from a single zygote (conceptus)

FETAL-MATERNAL CELL TRAFFICKING
- Goes both ways
- Carry with us for lifetime
- Stem cells called upon by injury/stress
- More cells cross when placenta in trouble
  - Bianchi, AJMG 2000; 91:22-28
  - Lo, AJMG 1999; 64:218-224
- Newborns have mother’s cells in many tissues
  - Bianchi, J Peds 2003; 142:31-35
- Relation to stem cell therapy

FETAL-MATERNAL CELL TRAFFICKING
- Goes both ways
- Carry with us for lifetime
- Stem cells called upon by injury/stress
- More cells cross when placenta in trouble
  - Bianchi, AJMG 2000; 91:22-28
  - Lo, AJMG 1999; 64:218-224
- Newborns have mother’s cells in many tissues
  - Bianchi, J Peds 2003; 142:31-35
- Relation to stem cell therapy

MOTHER’S CELLS IN FETUS
- 4 male babies who died for other reasons
- Maternal cells present in
  - Liver all++
  - Spleen all++
  - Thymus all++
  - Skin 2/4
  - Thyroid 1/2

FETAL-MATERNAL CELL TRAFFICKING
- Goes both ways
- Carry with us for lifetime
- Stem cells called upon by injury/stress
- More cells cross when placenta in trouble
  - Bianchi, AJMG 2000; 91:22-28
  - Lo, AJMG 1999; 64:218-224
- Newborns have mother’s cells in many tissues
  - Bianchi, J Peds 2003; 142:31-35
- Relation to stem cell therapy

NON-TRADITIONAL MECHANISMS
- Mosaicism
- Microchimerism
- Imprinting
- Epigenetics
- Microbiome
- Fetal programming

EFFORTS AT NON-INVASIVE PRENATAL DIAGNOSIS
- Can find fetal cells in mother’s blood
- However, they are from all her pregnancies
- Somehow these cells are tolerated long term

FETAL MATERNAL MICROCHIMERISM
CHIMERA
- An individual who is a mixture of cells from two or more genetically distinct populations of cells which have been derived from more than one zygote
MICROCHIMERISM
- The presence of a small number of cells derived from another individual (transfusion, transplantation, twin, etc.)
MOASIC
- An individual comprised of two or more genetically distinct populations of cells derived from a single zygote (conceptus)
**Fetal Cells in Mother**

- Fetal cells present in thyroid tissue (probably both male and female)
- Many types of thyroid disease (adenoma, Hashimoto’s carcinoma, multi-nodular goiter, thyroiditis)
- Fetal cells can integrate and contribute to repair
- Role of fetal cell depends on its origin (cell type)


**Summary of Microchimerism**

- Fetal maternal microchimerism is common and probably universal in mammals
- Pregnancy associated microchimerism is increased by complications of pregnancy
- These microchimeric progenitor/stem cells play a role in repair of many tissues
- May trigger or facilitate autoimmune reactions probably depending on the compatibility of the antigens inherited from father
- Transgenerational “memory” may be important in other ways
- Fetal maternal microchimerism may be the reason women live longer than men (fetal cells even found in mom’s brain)

**Non-Traditional Mechanisms** (revealed by studying congenital anomalies)

- Mosaicism
- Microchimerism
- Imprinting
- Epigenetics
- Microbiome
- Fetal programming

**Genomic imprinting**

- Germ line specific modification which produces functional difference in expression of genetic material depending on whether the genetic material is maternally or paternally derived

**Evidence of Genomic Imprinting**

- Pronuclear transplantation
- Triploid phenotypes
- Deletions and duplications depending on parent of origin (P of O)
- Uniparental disomies
- See changes in size, behaviour and survival

Genome imprinting implies there are specific pieces of genetic information that must come from mother and others that must come from father to have normal development

Genome imprinting involves both an imprintable gene and imprinting of that gene (i.e., turning off, associated with methylation)
Both associated with 15p –
- PW-paternal
- AS-maternal
Also both associated with 15UPD
- PW-paternal
- AS-maternal
2 separate critical regions and imprinting centres

Methylation of the imprinted gene is associated with non-expression
This was a key to looking at gene expression
Methylation requires a methyl group (role of folic acid)

WHY IS THERE IMPRINTING?
- Conflict regarding whose genome control
  - Father optimize fitness of offspring
  - Mother survivor
- Defense against transposable elements & virus
  - Primitive protection system that affects gene expression
- Rheostat response to environment
  - Fine tuning

Non-traditional mechanisms (revealed by studying congenital anomalies)
- Mosaicism
- Microchimerism
- Imprinting
- Epigenetics
- Microbiome
- Fetal programming

Now the challenge is to understand how genes are turned on and off
(i.e., control of gene expression)
- DNA methylation
- Histone acetylation
- Non-coding RNAs – small, micro, long, circular
- Telomeres
- What else?

Epigenetics
EPI = on or over – therefore modulates gene expression without modifying DNA sequence
1. Heritable states in gene function that are not explained by change in DNA sequence
2. Heritable and reversible modifications of chromatin which modulate chromosome/gene function and expression and thus the resulting phenotype

Epigenetics
EPI = on or over – therefore modulates gene expression without modifying DNA sequence
3. Meiotically and mitotically heritable changes in gene expression that are not coded in the DNA itself
4. Of interest because they are a window into control and regulation of gene expression and are relevant to tissue specificity and timing of expression

What “tells” the genes to turn “on” and “off”?
Who “tells” them to do it?!
FEATURES OF TRANSCRIPTIONALLY ACTIVE & INACTIVE CHROMATIN

<table>
<thead>
<tr>
<th>CHROMATIN</th>
<th>Transcriptionally ACTIVE chromatin</th>
<th>Transcriptionally INACTIVE chromatin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conformation</td>
<td>Open, extended conformation</td>
<td>Highly condensed conformation, particularly apparent in heterochromatin</td>
</tr>
<tr>
<td>DNA METHYLATION</td>
<td>Relatively unmethylated, especially at promoter regions</td>
<td>Methylated including at promoter regions</td>
</tr>
<tr>
<td>HISTONE ACYLATION</td>
<td>Acetylated histones</td>
<td>Deacetylated histones (e.g., dephosphorylation, etc.)</td>
</tr>
</tbody>
</table>

WHOLE SETS OF PROTEINS AND ENZYMES THAT MODIFY CHROMATIN

- Histone acetylases
- Histone deacetylases
- Histone methyltransferase
- Adenosine 5' di and triphosphate ribosylases
- DNMT (DNA methyltransferase) family
- MECP2 (time and tissue specific)

Epigenetic changes modulate gene expression without modifying DNA sequence
Epigenome is the total epigenetic changes at a particular time and in a particular tissue that will be present on top of the genome
NON-TRADITIONAL MECHANISMS
(revealed by studying congenital anomalies)
• Mosaicism
• Microchimerism
• Imprinting
• Epigenetics
• Microbiome
• Fetal programming

MICROBIOME
• Live in harmony
• Actually need them to be healthy
• 27 different zones
• Present prenatally
• Get new ones at the time of birth
• More than 10x as many bugs as cells

CHANGES OVER AGING
• In order to survive, mammals developed capacity for flexibility (plasticity)
• The evolution of mammals (humans) involved developing a spectrum of responses, survival advantage was to have flexibility in responses to a change in the environment
• Fetus particularly responsive and sensitive to maternal messages (potential mismatch)
• What are the types of signals from mom to fetus?
• Maternal diet seems to be a signal to the fetus, what else?
• Transgenerational effects

NON-TRADITIONAL MECHANISMS
(revealed by studying congenital anomalies)
• Mosaicism
• Microchimerism
• Imprinting
• Epigenetics
• Microbiome
• Fetal programming

FETAL PROGRAMING
(Development Origins of Health and Disease)
• Mammals evolved protective mechanisms
• Humans are mammals
• Transgenerational effects — many tissues, at many times, and gender specific effects
• Related to regulating growth/size, behavior, and survival

BARKER HYPOTHESIS
(Developmental Origins of Adult Health and Disease – DOHaD)
• Epidemiologist in UK
• Looking for cause of early onset cardiovascular disease
• Males with heart attacks at less than 40 years old
• Had intrauterine growth restriction and low socioeconomic status

CRITICAL/SENSITIVE PERIODS
• Barker Hypothesis:
  Critical/sensitive periods in development during which insult or stimulus causes long lasting effects on structure and/or function.
• Nutrition programs the structures and functions of the developing fetal organs resulting in susceptibilities which interact with environment to develop diseases later in life
Barker Hypothesis

- Undernutrition at different stages of pregnancy, and infancy leads to increased risk of:
  - Cardiovascular disease
  - Hypertension
  - Diabetes mellitus
  - Abnormal cholesterol levels
  - Abnormal coagulation factors
  - Osteoporosis
  - Cancer

(MUST BE TAKEN INTO ACCOUNT FOR GENOMIC ASSOCIATION STUDIES OF COMPLEX DISORDERS)

Fetal Origins of Adult Disease – Birth Size is Important

- Animal studies
- WWII blockade (Dutch famine)
- Immigrants
- Scandinavian studies of famine
- Quebec ice storm
- Differences in male and female offspring
- Many human populations

These are 3 generation effects

Fetal Programming

(Development Origins of Health and Disease)

- Evolutionary protective mechanism for mammals
- Humans are part of a mammalian species
- Transgenerational effects – probably several modes of communication – many organ specific effects
- Related to down regulating growth/size and behaviour
- Provides Darwinian fitness advantage adaptive responses
- Has potential for “mismatch” with later environment

“What gets inherited is not a deterministic genotype, but rather a genotype that encodes a potential range of phenotypes.”

Gilbert 2000

Genes are expressed in...

- Time in development specific
- Tissue specific
- Gender specific
- Parent of origin specific
- Pathways
- With transgenerational effects

And can be identified as silent or expressing

Is this complex or what?

Did we really think it would be simple?

Were you paying attention when all of this was happening?

Well – your tissues were!!
<table>
<thead>
<tr>
<th>AMC</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence</td>
<td>2/1000</td>
</tr>
<tr>
<td>Variability</td>
<td>&gt; 400 types</td>
</tr>
<tr>
<td>Stability</td>
<td>2/1</td>
</tr>
<tr>
<td>Prenatal Dx</td>
<td>Need gene</td>
</tr>
<tr>
<td>IUGR</td>
<td>With infection</td>
</tr>
<tr>
<td>PT/OT</td>
<td>Essential mobilization</td>
</tr>
<tr>
<td>Drugs Help</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Sometimes it seems a miracle that anyone survives when you know all the things that can go wrong**

**Principle of clinical/medical/human genetics is to utilize the unusual/exceptional case to learn general principles of biology**

Mother Nature is very clever, she tests things out, then reuses “good” pathways in additional ways, and she has built in the ability to adjust to different environments.

Our challenge is to work with Mother Nature to achieve optimum health for all of us, because nobody is “normal”!

"...that anyone survives when you know all the things that can go wrong..."
**Diagnostic Clinical Genome and Exome Sequencing**

Leslie G. Biesecker, M.D., and Robert C. Green, M.D., M.P.H.

**S**equencing of the genome or exome for clinical applications, hereafter referred to as clinical genome and exome sequencing (CGES), has now entered medical practice. Several thousand CGES tests have already been ordered for patients, with the goal of establishing diagnoses for rare, clinically unrecognizable, or puzzling disorders that are suspected to be genetic in origin. We anticipate increases in the use of CGES, the key attribute of which — its breadth — distinguishes it from other forms of laboratory testing. The interrogation of variation in about 20,000 genes simultaneously can be a powerful and effective diagnostic method.

CGES has been hailed as an important tool in the implementation of predictive and individualized medicine, and there is intense research interest in the clinical benefits and risks of sequencing for screening healthy persons; however, current practice recommendations do not support the use of sequencing for this purpose, and for that reason we do not further address it here. We have also limited this overview of CGES to the analysis of germline sequence variants for diagnostic purposes and do not discuss the use of CGES to uncover somatic variants in cancer in order to individualize cancer therapy.

Clinicians should understand the diagnostic indications for CGES so that they can effectively deploy it in their practices. Because the success rate of CGES for the identification of a causative variant is approximately 25%, it is important to understand the basis of this testing and how to select the patients most likely to benefit from it. Here, we summarize the technologies underlying CGES and offer our insights into how clinicians should order such testing, interpret the results, and communicate the results to their patients (an interactive graphic giving an overview of the process is available with the full text of this article at NEJM.org).

**TECHNICAL OVERVIEW AND LIMITATIONS OF CGES**

Detailed technical descriptions of sequencing can be found elsewhere, and we provide a graphical summary of one method of CGES in Figures 1 and 2. Regardless of the specific technology that is used, the process begins with the extraction of DNA from white cells, after which the DNA is broken into short fragments, the sequences of which are determined with the use of one of various sequencing technologies. The sequencing instrument generates millions of short sequence reads, which are strings of data representing the order of the DNA nucleotides, or bases, in each fragment. These sequence reads are then aligned to specific positions in the human genome reference sequence (see Glossary) with the use of computers. Similarities and differences between the patient’s sequence and the reference sequence are tabulated, and a computerized determination of the specific genotype (A, C, G, or T) at each position in the exome or genome is performed, resulting in an output file along
with information representing the number of sequence reads generated (depth of coverage) and the accuracy of the genotype at each position. The output file is computationally filtered in accordance with the clinical objective of the test and the preferences of the laboratory. Typically, the file is filtered for variants that are rare or have not previously been reported (because it is reasoned that a common variant cannot cause a rare disease), variants predicted to cause a loss or altered function of a gene, and variants previously reported to cause disease.10,11

CGES is most useful for the detection of single-nucleotide substitutions and insertions or deletions of 8 to 10 nucleotides or smaller; it is less accurate for other types of genomic variation (Table 1). The yield of sequence reads is inherently uneven across the exome (or genome) — typical results provide adequate coverage of 85 to 95% of the targeted sequence. With exome sequencing, there is also variable coverage of flanking intronic regions, which may include disease-causing variants that affect the splicing of messenger RNA encoded by the gene (splice variants).

**Glossary**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exome sequencing</td>
<td>DNA sequencing that targets the exons of all genes in the genome. The exome makes up about 1% of the genome, primarily exons of genes that code for proteins. This type of sequencing is sometimes referred to as “whole-exome sequencing,” even though coverage of the exons is not 100%.</td>
</tr>
<tr>
<td>Exons</td>
<td>Segments of genes that are spliced together after gene transcription to form messenger RNA, which, in turn, is translated into protein.</td>
</tr>
<tr>
<td>Expressivity</td>
<td>Variation in the severity of a genetic disorder among persons with some features of the condition.</td>
</tr>
<tr>
<td>Filtering analysis</td>
<td>The process of excluding DNA variants from further consideration because of various attributes, with the use of bioinformatics and manual curation. For example, most filtering analyses exclude synonymous variants (DNA variants that are predicted not to change the amino acid sequence of a protein).</td>
</tr>
<tr>
<td>Genome sequencing</td>
<td>DNA sequencing that targets the entire genome. It is sometimes termed “genome shotgun sequencing” or “whole-genome sequencing,” even though coverage is not 100%.</td>
</tr>
<tr>
<td>Germline variant</td>
<td>A DNA sequence variant that was transmitted by means of a gamete (sperm or egg) or that was caused by a mutation in the zygote or at a very early stage of fetal development and is presumed to be present in all of a person’s nucleated cells.</td>
</tr>
<tr>
<td>Human genome reference sequence</td>
<td>A reference sequence that provides a haploid mosaic of different DNA sequences from multiple donors, which is revised periodically and is not necessarily normal.</td>
</tr>
<tr>
<td>Penetrance</td>
<td>The likelihood that a person with a causative variant in a gene has any recognizable symptom, sign, or laboratory feature of the disease associated with that variant.</td>
</tr>
<tr>
<td>Sanger sequencing</td>
<td>A method of sequence determination, invented by Frederick Sanger, that uses dideoxy terminator nucleotide chemistry, with the reaction products separated by gel electrophoresis.</td>
</tr>
<tr>
<td>Variant</td>
<td>A difference in a DNA sequence in comparison with the normal reference sequence. A variant may be benign (sometimes referred to as a polymorphism) or pathogenic (sometimes referred to as a mutation).</td>
</tr>
</tbody>
</table>

**Indications for Ordering CGES**

CGES is currently indicated for the detection of rare variants in patients with a phenotype suspected to be due to a mendelian (single-gene) genetic disorder, after known single-gene candidates have been eliminated from consideration or when a multigene testing approach is prohibitively expensive. Patients can be of any age but are commonly children, since many genetic conditions are manifested in childhood; evaluations are performed because parents are searching for the cause or for information to guide management and treatment and desire accurate information regarding the risk of recurrence, as noted below.

The preparation for ordering CGES should include four key elements: gathering information on family history, systematically evaluating the patient’s phenotype, searching medical literature and databases, and obtaining informed consent. A thorough family history should be obtained to assess whether there are similar or related phenotypes in other family members, as well as to evaluate and assess the inheritance pattern. The patient (and other apparently affected family members) should be evaluated for other potentially relevant manifestations. For example, if the primary presentation is autism and CGES is being considered by a neurologist, the patient should also be carefully examined for respiratory, cardiac, renal, skin, and dysmorphic abnormalities. With a family history and a comprehensive phenotype in hand, a literature review or syndrome database search should be performed (Table 2) to determine whether the patient’s presentation matches a rare but estab-
Figure 1 (facing page). Schematic Overview of Exome Sequencing.
Exome sequencing targets the approximately 1% of the genome that is made up of exons, which encode protein sequence. The DNA from the patient (Panel A) is isolated and broken into fragments (Panel B); the DNA fragments are coupled to artificial DNA linker segments (Panel C), and the fragments are selected with the use of artificial DNA or RNA baits that are complementary to targeted DNA (not shown). The sequencing process starts with the binding of the end of each DNA fragment to a solid matrix and in situ amplification (Panel D), and the DNA fragments are then sequenced on the slide in a series of reactions in which a complementary nucleotide with one of four colored fluorescent dyes is added to each cluster of identical molecules (Panel E). The identity of the colored fluorescent indicator of each cluster is imaged with a laser and a camera coupled to a microscope, the fluorescent indicator is removed, and the cycle is repeated to generate a nucleotide sequence read that is 75 to 150 nucleotides in length. The sequence reads are aligned to a reference DNA sequence (Panel F), and a genotype call for each position is made. In this example, most of the positions are homozygous reference sequence, but one position is called as heterozygous A/T. This figure illustrates one widely used sequencing technology, but it is not intended to endorse that technology over other methods.

Figure 2. Schematic Comparison of Exome and Genome Sequencing.
Panel A shows the targeted nature of exome sequencing, with sequence reads concentrated over the coding portions of genes. This is in contrast to genome sequencing, shown in Panel B, in which the sequence reads are nearly randomly distributed over the entire genome. Each approach has advantages over the other, some of which are listed in the two panels.
There is a spectrum of genomic variants, from nucleotide insertions and deletions of at least 8 to 10 bp through copy-number variants, that are less effectively assayed by current CGES technology.

Table 1. DNA Variant Types Currently Not Well Detected or Undetectable by Clinical Genome and Exome Sequencing (CGES), with Examples of Phenotypes Associated with Such Variants.

<table>
<thead>
<tr>
<th>Variant Type</th>
<th>Associated Phenotype or Phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetitive DNA, including trinucleotide repeats</td>
<td>Fragile X syndrome, Huntington’s disease</td>
</tr>
<tr>
<td>Copy-number variants</td>
<td>DiGeorge syndrome (22q11.2 deletion syndrome), Charcot–Marie– Tooth disease type 1A</td>
</tr>
<tr>
<td>Long insertion–deletion variants†</td>
<td>Resistance to human immunodeficiency virus infection</td>
</tr>
<tr>
<td>Structural variants</td>
<td>Chromosomal translocations associated with spontaneous abortions</td>
</tr>
<tr>
<td>Aneuploidy</td>
<td>Down’s syndrome, Turner’s syndrome</td>
</tr>
<tr>
<td>Epigenetic alterations</td>
<td>Prader–Willi syndrome, Beckwith–Wiedemann syndrome</td>
</tr>
</tbody>
</table>

† There is a spectrum of genomic variants, from nucleotide insertions and deletions of at least 8 to 10 bp through copy-number variants, that are less effectively assayed by current CGES technology.

The outcomes of CGES analysis vary widely. In some CGES reports, a single causative variant is asserted to be the likely cause of the disease, whereas in others, multiple candidate variants are identified that must then be evaluated by the ordering clinician or by consultants. In many cases, no plausible variants are identified.

The two main considerations for evaluating CGES results are their analytic validity and their clinical validity. Analytic validity is a measure of the likelihood that the patient actually has the particular genotype shown in the CGES results — that is, the accuracy of the test. Clinical validity, which is much more complicated and challenging to assess, is the determination that a particular disease is truly caused by variants in a particular gene and that the specific variant that has been detected is indeed pathogenic.

Positive CGES findings are highly accurate, but the false negative rate varies according to the genomic region. For this reason, CGES is not yet a substitute for targeted sequencing of suspected genes or gene panels that have been optimized for a particular condition. Most laboratories validate positive CGES results with well-established methods such as polymerase-chain-reaction amplification and Sanger sequencing. Confirmatory Sanger sequencing should be considered when any major medical intervention is being contemplated on the basis of a CGES result, if such confirmation is not routinely provided by the CGES laboratory.

As noted above, determining the clinical validity of CGES results is more challenging than determining their analytic validity. The general approach is to compare variants implicated by CGES with databases of known variation, which are in turn based on reports that describe causal variants as well as associations between variants.
and phenotypes. In the literature, however, there are many false attributions of disease to variants, a problem that is in part due to the conflation of association with causation. Clinicians reviewing the results of sequencing should be aware of the possibility of a false attribution of pathogenicity to a variant and should realize that the chances of false attribution are increased in CGES because thousands of genes are tested simultaneously.

The clinical usefulness of identifying the variant that is the cause of a previously undiagnosed syndrome or heritable disorder varies. In some cases, it can lead to a specific treatment or management strategy that dramatically changes the clinical outcome. In the majority of cases in which the finding does not change clinical management, treatment, or prognosis, it may still be useful because it can end an expensive, potentially invasive, and stressful diagnostic odyssey. The identification of the causative variant may provide accurate estimates of recurrence risk and facilitate preconception intervention or prenatal diagnosis for the affected patient or affected at-risk relatives. In adult-onset disease, one of the most useful outcomes of successfully identifying the causative variant is the subsequent detection of presymptomatic, at-risk siblings for whom screening or preventive therapy might improve the clinical outcome. Examples include enhanced surveillance or prophylactic surgery for patients found to have a genetic susceptibility to cancer.

Pretest counseling is particularly important, to maintain realistic expectations for finding the causative variant and to alert the patient or family that in most cases, a positive result is unlikely to change treatment or management decisions or to improve the prognosis. In addition, the patient should be advised that incidental findings unrelated to the reason for testing may be found and reported, as described below. It may also be important to discuss the cost of the test with the patient. As is the case with many medical services, assessment of the cost is complicated by many factors. The published billing charge for CGES in most laboratories is in the range of $4,000 to $15,000 per patient, with some laboratories offering lower per-person charges for family testing. To put this in perspective, the per-person charge for sequencing of an exome may be only two to four times the published billing charge for some single-gene sequencing tests, which is why exome sequencing can be more efficient in a number of clinical scenarios. Some laboratories have reported that third-party payers are reimbursing for this testing, but practices vary widely, and patients should understand this in advance.

**Table 2. Examples of Online Databases to Assist Clinicians in Differential Diagnosis or Candidate-Gene Identification for Rare Syndromic Disorders before CGES Is Performed.**

<table>
<thead>
<tr>
<th>Table 2. Examples of Online Databases to Assist Clinicians in Differential Diagnosis or Candidate-Gene Identification for Rare Syndromic Disorders before CGES Is Performed.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Free access (or an available free-access version)</strong></td>
</tr>
<tr>
<td>HuGE Navigator (<a href="http://hugenavigator.net/HuGENavigator">http://hugenavigator.net/HuGENavigator</a>)</td>
</tr>
<tr>
<td>Human Gene Mutation Database (<a href="http://www.biobase-international.com/product/hgmd">www.biobase-international.com/product/hgmd</a>)</td>
</tr>
<tr>
<td>Online Mendelian Inheritance in Man (<a href="http://www.omim.org">www.omim.org</a>)</td>
</tr>
<tr>
<td>Phenomizer (<a href="http://compbio.charite.de/phenomizer">http://compbio.charite.de/phenomizer</a>)</td>
</tr>
<tr>
<td>SimulConsult (<a href="http://www.simulconsult.com">www.simulconsult.com</a>)</td>
</tr>
<tr>
<td><strong>Subscription or fee required for access</strong></td>
</tr>
<tr>
<td>Isabel (<a href="http://www.isabelhealthcare.com/home/default">www.isabelhealthcare.com/home/default</a>)</td>
</tr>
<tr>
<td>London Medical Databases (<a href="http://lmdatabases.com">http://lmdatabases.com</a>)</td>
</tr>
<tr>
<td>POSSUM (<a href="http://www.possum.net.au">www.possum.net.au</a>)</td>
</tr>
</tbody>
</table>

**Interpreting and Communicating CGES Results**

Clinicians should review the CGES results delivered by the laboratory geneticist and place the findings into context with other relevant medical considerations. Sometimes an identified variant will spur additional history taking or an additional examination of the patient, which may reveal clinical features of a previously unrecognized syndrome or lead to the conclusion that the variant is not related to the disorder in the patient (Table 4).

In some cases, the CGES report from the testing laboratory identifies a causative variant (or two variants for a recessive disorder) in a single gene that is considered sufficiently pathogenic and specific that a diagnostic association with a heritable disorder is strongly supported. Such a conclusion by the laboratory geneticist is typically based on the integration of the submitted clinical information with information on diseases associated with the identified variant. In this case, as with all laboratory tests, the order-
Table 3. Genetic Considerations in Deciding Which Relatives Should Undergo CGES.

<table>
<thead>
<tr>
<th>Autosomal dominant inheritance suspected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing is most valuable in the trio consisting of a child and both biologic parents if the disorder may be caused by a de novo variant of a gene associated with a disorder that is caused by heterozygous (monoallelic) mutations.</td>
</tr>
<tr>
<td>If a phenotype with autosomal dominant inheritance within a family is being evaluated, sequences from two distant relatives with the phenotype will be more valuable than sequences from two close relatives with the phenotype. Sequences from distant relatives are recommended because there will be fewer variants shared solely by chance in close relatives.</td>
</tr>
<tr>
<td>Autosomal recessive inheritance suspected</td>
</tr>
<tr>
<td>Testing can also be valuable in the trio of a child and both biologic parents if a gene variant inherited in an autosomal recessive pattern are being considered.</td>
</tr>
<tr>
<td>If there is consanguinity, then testing the trio may be less helpful, because the child’s sequence will already have large areas of homozygosity.</td>
</tr>
</tbody>
</table>

Table 4. Examples of How to Evaluate the Pathogenicity of a Variant.

<table>
<thead>
<tr>
<th>Variant supported as pathogenic</th>
</tr>
</thead>
<tbody>
<tr>
<td>A CGES report for a boy with apparently isolated intellectual disability identified a novel, previously unidentified variant in the gene ATRX. This gene is associated with the α-thalassemia mental retardation syndrome. In response to this report, the clinician orders hematologic testing, which identifies a subtle thalassemia phenotype. This additional testing strongly supports the variant identified in the CGES result as the cause of the intellectual disability in the patient.</td>
</tr>
<tr>
<td>Variant not supported as pathogenic</td>
</tr>
<tr>
<td>A CGES report identified a variant in the gene NF2 in a 2-month-old infant with congenital, bilateral sensorineural hearing loss. The variant is present in the Human Gene Mutation Database as a disease-causing variant associated with neurofibromatosis 2. However, that database entry is based on a single report that did not specify whether the patient with the variant was a case patient, a patient with a suspected case, or a control. A review of neurofibromatosis 2 in GeneReviews (<a href="http://www.ncbi.nlm.nih.gov/books/NCBI1201/">www.ncbi.nlm.nih.gov/books/NCBI1201/</a>) shows that this disorder typically causes unilateral hearing loss with an onset in young adulthood, not bilateral deafness with an onset in infancy. This post-hoc assessment — the absence of support in the literature for causality of the variant combined with data from GeneReviews on the clinical features known to be caused by mutations in NF2 — suggests that the evidence for the pathogenicity of this variant is weak, and the variant is unlikely to explain the child’s phenotype.</td>
</tr>
</tbody>
</table>

CGES can generate results unrelated to the indication that prompted sequencing, and these results may be clinically useful. As with all testing, it is important to constrain such evaluations to avoid unnecessary future testing and expense. To this end, the American College of Medical Genetics and Genomics has recommended that the laboratories providing CGES routinely seek and report to the ordering clinician specific variants in a minimum set of 56 referring to a specialist, depending on the nature of the diagnosed disorder.

In other cases, the laboratory CGES report identifies one or more candidate variants that may or may not be the cause of the phenotype that triggered ordering of the test. Additional literature and database searching, additional phenotyping, and even functional studies may be needed to winnow these candidates down to the causative variant, if indeed it is among these candidates. This approach to post-test clinical evaluation requires a high level of expertise in genetics and informatics, as well as knowledge of the clinical features of the phenotype in the patient and of the phenotypes associated with the identified variants and genes. The details of this type of evaluation are beyond the scope of this review, but such evaluations may require consultation with geneticists or other disease-specific experts and the enrollment of patients and their families in studies.

If a CGES result is negative, subsequent improvements in knowledge may lead to the recognition that a previously uninterpretable variant in a negative CGES result is in fact pathogenic. The methods and approaches for ongoing reanalysis of CGES results have not been established, but it should eventually be possible to regularly reanalyze such results with the goal of identifying previously unknown variants. This will extend the scope of this review, but such evaluations may require consultation with geneticists or other disease-specific experts and the enrollment of patients and their families in studies.
CGES is a useful diagnostic test for a number of clinical situations, and it is already being used by clinical geneticists and other specialists. The indications and approaches we outline here are sure to evolve over time, as more data are generated for various clinical disorders, data interpretation is improved, and CGES is studied in new clinical situations (e.g., in neonatal medicine). Clinicians can feel confident ordering this test if they become comfortable with the evaluations, both before and after testing, and the processes that are required to maximize its usefulness, and if they are familiar with its limitations.

The opinions expressed in this article reflect the views of the authors and may not represent the opinions or views of any institutions with which they are affiliated.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the members of the Clinical Sequencing Exploratory Research Consortium, for valuable ideas and discussion on this topic, and Shamil Sunyaev, Ph.D., for helpful suggestions.

**REFERENCES**